PCT/GB 98 / 0 1 2 3 3

09/423037

The Patent Office

The Patent Office
Cardiff Road
Newport
Gwent
NP9 1RH

RECD 16 JUN 1998
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation and Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

PRIORITY DOCUMENT

Signed A Bennett.

Dated 18 May 1998

BEST AVAILABLE COPY

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

J
30 APR 1997
RECEIVED BY POST

30APR97 E271155-2 002934
P01/7700 25.00 - 9708676.3

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference PHM. 97/021

2. Patent application number *(The Patent Office will fill in this part)* **9708676.3** 30 APR 1997

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*
Imperial Cancer Research Technology Limited
Sardinia House
Sardinia Street, London, WC2A 3NL, United Kingdom
Patents ADP number *(if you know it)* ~~6254007002~~ 41136190001
If the applicant is a corporate body, give the country/state of its incorporation

4. Title of the invention
CHEMICAL COMPOUNDS

5. Name of your agent *(if you have one)*
GILES, Allen Frank
"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*
Intellectual Property Department
ZENECA Pharmaceuticals
Mereside, Alderley Park
Macclesfield, Cheshire, SK10 4TG, United Kingdom
Patents ADP number *(if you know it)* 6988463001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

9. Enter the number of sheets for any [redacted]e following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	17
Claim(s)	–
Abstract	–
Drawing(s)	6 + 6

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature A F Giles Date 24-APR-97

12. Name and daytime telephone number of person to contact in the United Kingdom

ALLEN F GILES
01625 516573

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

CHEMICAL COMPOUNDS

The present invention relates to inhibition of the interaction between nuclear proteins and nuclear receptors through identification of the key structural element responsible for the interaction.

The binding of lipophilic hormones, retinoids and vitamins to members of the nuclear receptor (NR) superfamily (to form "liganded" receptors) modifies their DNA binding and transcriptional properties, resulting in the activation or repression of target genes [1,2]. Ligand binding induces conformational changes in NRs and promotes their association with a diverse group of nuclear proteins, including SRC-1/p160 [3,4,5], TIF2 [6,7] and CBP/p300 [4,5,8,9] which function as coactivators, and RIP-140 [10], TIF1 [11] and TRIP1/SUG1 [12, 13] whose functions are unclear.

The recruitment of nuclear proteins (coactivators and/or other so-called bridging proteins) by NRs is thought to be essential to their function as ligand-induced transcription factors. Structural studies of the ligand binding domains (LBDs) of three different nuclear hormone receptors, the retinoid X receptor α (RXRα) [15], the retinoic acid receptor γ (RARγ) [16] and the thyroid hormone receptor β (TRβ) [17], have led to the proposal that binding of ligand results in a realignment of a conserved amphipathic α-helix, Helix 12 (H12), generating a novel surface required for coactivator binding and consequently activator function 2 (AF2)-dependent transactivation. Consistent with this, mutations of conserved hydrophobic residues in H12 which impair AF2 [14, 18-20], also interfere with the ability of NRs to bind coactivators [4,6,10,11,13]. Less is known about the coactivator sequences which mediate interaction with NRs although several proteins appear to contain multiple NR binding sites [5,8,21]. However the present state of knowledge is completely silent about precisely how liganded nuclear receptors interact with nuclear proteins to modify their DNA binding and transcriptional properties, resulting in the activation or repression of target genes.

The present invention is based on the discovery that a short signature motif present in the nuclear proteins is necessary and sufficient to mediate their binding to liganded NRs.

According to one aspect of the present invention there is provided a method of reducing the interaction between

a) a first region which is a signature motif on a nuclear protein, and

b) a second region which is that part of a nuclear receptor which is capable of interacting with the nuclear protein through binding to the signature motif,

in which the method comprises adding an inhibitor in the presence of the nuclear receptor and the nuclear protein, the inhibitor being characterised in that it reduces the interaction between the first region of the nuclear protein and the second region of the nuclear receptor.

The term "nuclear protein" means the bridging factors (including coactivators) that are responsible for the interaction between a liganded nuclear receptor and the transcription initiation complex involved in regulation of gene expression (reviewed for steroid hormone receptors in Beato, M., Herrlich, P. & Schutz, G. *Cell* **83**, 851-857 (1995)). The term "nuclear receptor" means the family of nuclear receptors such as described in Mangelsdorf, D.J., *et al. Cell* **83**, 835-839 (1995). The term "signature motif" means a short sequence of 4-10 amino acid residues which is the key structural element of a nuclear protein which binds to a liganded nuclear receptor as part of the process of the activation or repression of target genes.

Preferably the signature motif is B1XXLL in which B1 is any natural hydrophobic amino acid, L is leucine and X represents any natural amino acid. Preferably B1 is leucine or valine with leucine being most preferred. In some instances the preferred signature motif is further defined as B2B1XXLL wherein "B2" is a hydrophobic amino acid residue as defined for B1. A "natural hydrophobic amino acid" is defined as any one of isoleucine, leucine, methionine, phenylalanine, tryptophan, tyrosine or valine. Preferably the signature motif is in the conformation of a helix and the leucine residues form a hydrophobic face thereof. Preferably the signature motif is positioned within a molecule so that it is available at the surface thereof for interaction with proteins.

Herein we show that the ability of a nuclear protein (SRC1) to bind a nuclear receptor (liganded ER) and enhance its transcriptional activity is dependent upon the integrity within the nuclear protein of the signature motif (LXXLL), as well as key hydrophobic residues in the conserved helix (Helix 12) of NRs required for their ligand-induced activation function (AF-2) [14]. The signature motif is also found in TIF1, TIF2, p300, RIP 140 and the TRIP proteins, and occurs within regions of these proteins known to be sufficient for interaction with NRs. Thus the LXXLL motif is a signature sequence which facilitates the interaction of

diverse proteins with nuclear receptors, and thus is a key part of a new family of nuclear proteins.

A preferred nuclear protein is a coactivator, in particular the nuclear protein includes any one of RIP 140, SRC-1, TIF2, CBP, p300, TIF1, Trip1, Trip2, Trip3, Trip4, Trip5, Trip8 or Trip9.

In this specification a reference to a nuclear protein or nuclear receptor includes isoforms thereof unless stated or otherwise implicit from the context. An isoform is a family of functionally related proteins that differ slightly in their amino acid sequences such as for example from differential splicing of exons during transcription. SRC1a is an example of an isoform of SRC1.

Nuclear receptors are transcription factors. A preferred transcription factor comprises at least part of a conserved amphipathic α-helix, and especially preferred is retinoic acid receptor or a steroid hormone receptor. Preferred steroid hormone receptors are oestrogen receptor, progesterone receptor, androgen receptor and glucocorticoid receptor with oestrogen receptor being especially preferred.

Preferably the second region comprises at least part of a conserved amphipathic α-helix such as for example Helix 12 in the oestrogen receptor which is especially preferred.

An especially preferred combination of nuclear receptor and nuclear protein is one in which the nuclear receptor is oestrogen receptor and the nuclear protein is selected from SRC1, TIF2, CBP and p300, with SRC1 and especially SRC1a being most preferred.

According to another aspect of the present invention there is provided a novel inhibitor as described above. Preferably the inhibitor is a peptide, more preferably a peptide comprising the signature motif defined above, and more preferably the peptide has less than 15 amino acid residues. Especially preferred inhibitors are any one of the following peptides; PQAQQKSLLQQLLT, KLVQLLTTT, ILHRLLQE, or LLQQLLTE. Peptides may be prepared using conventional techniques for example using solid phase synthesis and Fmoc chemistry. These peptides are useful in the treatment of oestrogen responsive tumours.

A novel inhibitor, for example, could be an antibody against a signature motif or a novel small molecule which binds to the signature motif or its complementary binding target (nuclear receptor second region) such that normal biological activity is prevented. Examples of small molecules include but are not limited to small peptides or peptide-like molecules.

The term "antibodies" is meant to include polyclonal antibodies, monoclonal antibodies, and the various types of antibody constructs such as for example $F(ab')_2$, Fab and single chain Fv. Antibodies are defined to be specifically binding if they bind with a K_a of greater than or equal to about 10^7 M^{-1}. Affinity of binding can be determined using conventional techniques, for example those described by Scatchard *et al.*, *Ann. N.Y. Acad. Sci.*, 51: 660 (1949).

Polyclonal antibodies can be readily generated from a variety of sources, for example, horses, cows, goats, sheep, dogs, chickens, rabbits, mice or rats, using procedures that are well-known in the art. In general, immunogen is administered to the host animal typically through parenteral injection. The immunogenicity may be enhanced through the use of an adjuvant, for example, Freund's complete or incomplete adjuvant. Following booster immunizations, small samples of serum are collected and tested for reactivity. Examples of various assays useful for such determination include those described in: *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988; as well as procedures such as countercurrent immuno-electrophoresis (CIEP), radioimmunoassay, radioimmunoprecipitation, enzyme-linked immuno-sorbent assays (ELISA), dot blot assays, and sandwich assays, see U.S. Patent Nos. 4,376,110 and 4,486,530.

Monoclonal antibodies may be readily prepared using well-known procedures, see for example, the procedures described in U.S. Patent Nos. 4,902,614, 4,543,439 and 4,411,993; Monoclonal Antibodies, Hybridomas: *A New Dimension in Biological Analyses,* Plenum Press, Kennett, McKearn, and Bechtol (eds.), (1980).

Monoclonal antibodies can be produced using alternative techniques, such as those described by Alting-Mees et al., "Monoclonal Antibody Expression Libraries: A Rapid Alternative to Hybridomas", *Strategies in Molecular Biology* 3: 1-9 (1990) which is incorporated herein by reference. Similarly, binding partners can be constructed using recombinant DNA techniques to incorporate the variable regions of a gene that encodes a specific binding antibody. Such a technique is described in Larrick *et al.*, *Biotechnology*, 7: 394 (1989).

According to a further feature of the invention there is provided a pharmaceutical composition which comprises a novel inhibitor of the invention, or a pharmaceutically-acceptable salt thereof, in association with a pharmaceutically-acceptable diluent or carrier.

The composition may be in a form suitable for oral use, for example a tablet, capsule, aqueous or oily solution, suspension or emulsion; for topical use, for example a cream, ointment, gel or aqueous or oily solution or suspension; for nasal use, for example a snuff, nasal spray or nasal drops; for vaginal or rectal use, for example a suppository; for administration by inhalation, for example as a finely divided powder such as a dry powder, a microcrystalline form or a liquid aerosol; for sub-lingual or buccal use, for example a tablet or capsule; or for parenteral use (including intravenous, subcutaneous, intramuscular, intravascular or infusion), for example a sterile aqueous or oily solution or suspension. In general the above compositions may be prepared in a conventional manner using conventional excipients. For peptidic inhibitors, parenteral compositions are preferred.

The amount of active ingredient that is combined with one or more excipients to produce a single dosage form will necessarily vary depending upon the host treated and the particular route of administration. For example, a formulation intended for oral administration to humans will generally contain, for example, from 0.5 mg to 2 g of active agent compounded with an appropriate and convenient amount of excipients which may vary from about 5 to about 98 percent by weight of the total composition. Dosage unit forms will generally contain about 1 mg to about 500 mg of an active ingredient.

According to another aspect of the present invention there is provided a method of mapping nuclear receptor interaction domains in nuclear proteins in which the method comprises analysis of the sequence of a nuclear protein for the presence of signature motifs as defined herein in order to identify an interaction domain. Preferably the analysis further comprises analysis of any potential interaction domains identified thereby for α-helicity and/or surface accessibility.

The invention is illustrated by the non-limiting Examples below in which, unless stated otherwise: temperatures are expressed in degrees Celsius; and sequences are listed N-terminus to C-terminus.

Figures 1a/1b show the interaction of LXXLL motifs derived from coactivators with the ER.

Figure 1a: Yeast two hybrid interactions of LXXLL motifs, derived from the proteins RIP140, SRC1a and CBP with the LBDs of wild type or mutant ER. The sequences of the LXXLL motifs in the DNA binding domain (DBD) fusion proteins are indicated. DBD-

LXXLL proteins were coexpressed with AAD-ER or AAD-ER Mut, which consist of an acidic activation domain (AAD) fused to the LBD of the wild-type ER, or a transcriptionally defective ER mutant, respectively. Reporter activities were determined in the presence or absence of 10^{-7}M 17-β-estradiol (E2) and expressed as units of β-galactosidase activity.

Figure 1b: Effects of mutations in the RIP140 LXXLL motif located at amino acids 935-943 on binding of AAD-ER. Conserved leucine residues are boxed and mutated residues are circled. The reporter activity was determined in the presence (black bars) or absence (white bars) of 10^{-7}M E2.

Figures 2a/2b/2c show that LXXLL motifs are required for binding of SRC1 to the ER LBD *in vitro* and for the ability of SRC1 to enhance ER activity *in vivo*.

Figure 2a: Wild type (SRC1a) and mutant (SRC1a-M1234) SRC1 proteins are shown schematically. The black bars represent the approximate locations of the LXXLL binding motifs in the linear SRC1a sequence and the shaded circles indicate the mutation of LXXLL binding motifs by replacement of conserved leucine residues with alanines (see Methods). Binding of wild type SRC1a or SRC1a-M1234 mutant to glutathione S transferase (GST) alone, to the ligand binding domain (aa 313-599) of ER (GST-AF2), or the SRC1 binding domain (aa 2058-2163) of CBP (GST-CBP) in the presence (+) and absence (-) of 10^{-6}M E2. The signals obtained with 10% of the input of [^{35}S] -labelled wild type and mutant SRC1 proteins are shown.

Figure 2b: The ability of increasing amounts of the peptides P-1 and P-2 to compete against the binding of wild type SRC1a to GST-AF2 in the presence of ligand is shown. The sequences of the P-1 and P-2 peptides are given at the foot of Fig 2b, and the conserved leucines and alanine substitutions are boxed.

Figure 2c: Wild type but not mutant SRC1e M123 potentiates activation by ER of the reporter gene 2ERE-pS2-CAT in transiently transfected Hela cells. Reporter activities obtained from extracts of transfected cells grown in the absence (white columns) or presence (black columns) of ligand (10^{-8}M E2). The amounts of ER, SRC1-wt and and SRC1-mut expression plasmids used in the transfections are indicated below the graph. The activities shown are averaged from duplicates.

Figures 3a/3b show that the LXXLL sequence is a signature motif in proteins that bind the LBDs of NRs.

Figure 3a: Alignment of LXXLL motif sequences present in human RIP140 [10], human SRC1a, mouse TIF2 [6], mouse CBP [23,24], p300 [33], mouse TIF1 [11] and human TRIP proteins [12]. The conserved leucines are boxed and the amino acid numbers are given for each motif.

Figure 3b: Schematic representation of the incidence of LXXLL motifs (black bars) in the sequences of proteins which bind NRs. The amino acid boundaries of the known NR binding sites are also shown.

The following abbreviations are used.

AAD	acidic activation domain
AF	activator function
DBD	DNA binding domain
E2	17-β-estradiol
ER	estrogen receptor
GST	glutathione S transferase
LBD	ligand binding domain
NR	nuclear receptor
PCR	polymerase chain reaction
RAR	retinoic acid receptor
RIP	receptor interacting protein
RXR	retinoid X receptor
SRC	steroid receptor coactivator
TIF	transcriptional intermediary factor
TRβ	thyroid hormone receptor β

Standard amino acid abbreviations have been used.

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N

Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Point mutations will be referred to as follows: natural amino acid (using the 1 letter nomenclature) , position, new amino acid. For example "L636A" means that at position 636 a leucine (L) has been changed to alanine (A). Multiple mutations will be shown between square brackets.

Example 1

Mapping of interaction sites between Nuclear Receptor and Nuclear Protein

It has been previously demonstrated that the 140 kDa receptor interacting protein (RIP140) bound directly to NRs through at least two distinct sites located at the N- and C-termini of the protein [21]. To map these interaction sites in more detail, we examined a series of twenty different PCR-generated fragments of RIP140 coding sequence fused in frame with a heterologous DBD, for interaction with NRs in a two hybrid system. Remarkably, although the different constructs spanned the entire 1158 amino acids of RIP140 sequence, all but two

displayed ligand-dependent interaction with ER, including five non-overlapping RIP140 sequences. By comparison of the sequences of the shortest interacting fragments we identified a short motif (LXXLL) common to all interacting fragments. In total, nine copies of the motif were identified in the RIP140 sequence, but the motif was absent in fragments showing no binding activity in our experiments.

To determine if these short sequences were sufficient to bind to NRs, we constructed a series of proteins consisting of a DBD fused to eight to ten amino acids incorporating one copy of each of the nine LXXLL motifs. As shown in **Fig.** 1a, each of the nine motifs present in RIP140 displayed strong ligand-dependent interaction with the LBD of the ER whereas the DBD alone showed no ability to bind. (Note that of the 10 motifs listed for RIP140, the 10th is a repeat of the 9th locus). Comparable results were obtained with the LBD of RAR (data not shown). Mutation of hydrophobic residues within H12 abolish AF2 activity and prevent the recruitment of RIP140 [10], TIF1 [11], TIF2 [6], SUG1 [13] and SRC1. Similarly, mutation of H12 residues M543 and L544 in the ER abolished the ligand-dependent interaction of all nine LXXLL motifs with ER (**Fig.** 1a). Taking these results together, we conclude that a short conserved motif comprised within as little as eight amino acids is sufficient to bind to transcriptionally active NRs. This discovery that such a relatively small motif can affect the interaction between two relatively large molecules is unprecedented in this field.

Secondary structure analysis using the Phd program [22] revealed that each of the nine copies of this motif in RIP140 occurred within a region predicted to be α-helical in nature, in which the conserved leucines would form a hydrophobic face.

Example 2

Mutational Analysis of a Signature Motif

To determine the sequence constraints required to observe a functional interaction, we carried out a partial mutational analysis of one of the RIP140 motifs (amino acids 935-943; **Fig.**1b). While western blot analysis showed no significant variation in the expression of the wild type and mutant fusion proteins (data not shown), mutation of valine 935 to alanine resulted in approximately ten fold reduction in the reporter activity in the presence of ligand which, when coupled with the observation that the first amino acid is hydrophobic in seven of the nine LXXLL motifs in RIP140, may indicate a preference for a hydrophobic residue at

this position. Strikingly, mutation of any one of the three conserved leucine residues L936, L939 or L940 to alanine resulted in a complete loss of binding to the LBD of ER (**Fig.** 1b) and RAR (data not shown), emphasising their importance in mediating the interaction with NRs. In contrast, mutation to alanine of L941 (which is not conserved among the motifs; see **Fig.** 3a), had no effect on the ability of this sequence to bind to the ER LBD. Replacement of a conserved leucine residue with a valine was tolerated at L936, but not at L939 or L940 indicating that hydrophobic character alone is not sufficient to maintain an interaction with ER (Fig. 1b). The amino acids K937, Q938, S942 and E943 were not subjected to mutagenesis as they are not conserved among the motifs we have identified (see Fig. 3a).

Example 3

Analysis of Signature Motifs in Nuclear Proteins

The steroid receptor coactivator SRC1, which stimulates ligand-dependent transcriptional activity, was originally identified as a partial cDNA encoding a protein capable of interacting with the progesterone receptor by means of a 196 amino acid C-terminal region [3]. We noted that the eight most C-terminal amino acids fit the LXXLL consensus, and indeed this sequence (DBD-SRC1a 1434-1441) displayed strong ligand-induced binding to ER, but not the ER H12 mutant (Fig. 1a). Subsequent studies have identified full-length SRC1 (SRC1a) from mouse (1459 amino acids) [4,5] and human (1441 amino acids) tissues. Both murine and human SRC1a proteins interact with multiple NRs, and contain an additional interaction region between residues 569 to 789 [5] and 570-780, respectively. Three copies of the LXXLL motif were identified in this central interaction domain of human SRC1a (see Fig. 3a & 3b), each of which displayed ligand-dependent binding to both ER (Fig. 1a) and RAR (data not shown) in the two hybrid assay, but not the ER H12 mutant. Interestingly, the sequences and relative positions of the three motifs in the central domain of SRC1a are conserved in the related coactivator protein Transcriptional Intermediary Factor 2 (TIF2) (Fig. 3a + b), and correspond to the region of TIF2 known to bind to NRs [6]. However, unlike SRC1a, TIF2 appears to lack a motif at its C-terminus. In addition, we noted that SRC1 contains three other sequences matching the LXXLL consensus. Although the motif at residues 45-53 is predicted by the Phd program to be α-helical and lies within the basic helix-loop-helix domain at the N-terminus of SRC1a [5], it showed only a very weak (6-fold)

interaction with liganded ER (Fig. 1a) or RAR (not shown) in the yeast two hybrid assay. This is consistent with the observed absence of strong NR-binding activity associated with the N-terminus of SRC1. The other two motifs within residues 111-118 and 912-920 both contain proline residues and are unlikely to adopt α-helical structure according to the Phd program. Indeed, these sequences showed no detectable interaction with NRs in our binding assays (Fig. 1a), which strongly suggests a preference for appropriate secondary structure for binding of LXXLL sequences to NRs.

Recent reports have indicated that CBP/p300 proteins, which were originally identified as coactivators for CREB [23,24], are coactivators for many transcription factors including NRs [4,5,8,9] and may serve as integrators of several signalling pathways [4]. CBP was shown to bind directly to NRs via its N-terminal 101 amino acids [4], with a possible RXR-specific binding site between residues 356-495 [8]. Our analysis showed that the CBP sequence harbours copies of the LXXLL motif within positions 68-78, and 356-364, which are conserved in the p300 sequence (amino acids 80-90 and 341-351; Fig 3a). Indeed, when tested in the two hybrid assay, the N-terminus of CBP (amino acids 1-101; data shown) and the LXXLL motif at residues 68-75 of the CBP sequence (Fig. 1a) displayed ligand-dependent binding to ER, but not the transciptionally defective ER mutant (Fig. 1a).

Example 4

The Binding Of Coactivator Proteins To NRs Is Dependent On Signature Motifs

To demonstrate that the binding of coactivator proteins to NRs is dependent on LXXLL motifs, we introduced alanine substitutions in SRC1a at the conserved leucine couplets at residues [L636A, L637A, L693A, L694A, L752A, L753A, L1438A, L1439A] thus effectively creating a mutant protein (SRC1a-M1234) in which all the four functional binding motifs were disabled. We then compared the ability of *in vitro* translated SRC1a and SRC1a-M1234 to bind to the ligand binding domain of the mouse estrogen receptor fused to glutathione-S-transferase (GST-AF2) in GST pulldown experiments (Maniatis et al.,(1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.). As shown in Fig. 2a, while wild type SRC1a protein displayed ligand-dependent binding to GST-AF2, SRC1a-M1234 failed to bind to GST-AF2 either in the presence or absence of ligand. To confirm that the mutations did not induce gross structural

disruption of the SRC1a-M1234, we compared the ability of the *in vitro* translated proteins tc interact with amino acids 2058-2163 of CBP, which was previously defined as the SRC1 binding domain [4]. Both proteins retained strong binding to GST-CBP (Fig. 2a) indicating that this SRC1 function remained intact in both wild type and mutant proteins. In addition we showed that the binding of wild type SRC1a to GST-AF2 was competed by increasing concentrations of a short peptide (P1) corresponding to the motif at the C-terminus of SRC1a (Fig. 2b). In contrast, a similar peptide (P2) in which the LXXLL motif was mutated, or peptides unrelated to the LXXLL motif (data not shown), did not compete the binding of SRC1a to GST-AF2 (Fig. 2b).

Finally, to demonstrate that LXXLL motifs are necessary for the function of SRC1 *in vivo,* we compared the abilities of wild type SRC1 and a mutant protein in which all LXXLL motifs were disabled to enhance the activity of mouse ER in transient transfection experiments. As shown in Fig. 2c, wild type SRC1 enhanced the activity of ER in a concentration-dependent manner. In contrast, the SRC1 mutant, which was unable to bind ER (Fig. 2a) had no stimulatory effect, but reduced ER activity by up to 50% at the highest concentration (Fig. 2c). This apparent dominant negative property of the mutant SRC1 is likely due to its ability to maintain interactions with CBP while failing to interact with NRs (Fig. 2a). This result is of interest given the recent evidence that SRC1 and CBP/p300 may exist as a complex *in vivo* [9], and that CBP also has NR binding activity [4, 8], as our data suggest that the interactions between NRs and CBP are sufficient to compensate for the inability of the SRC1a mutant protein to bind NRs, at least under these conditions. It remains to be determined whether NRs are engaged simultaneously by p160 and p300 proteins functioning independently or as a complex.

Examination of the sequences of other proteins known to bind to NRs revealed them to contain one or more copies of the LXXLL motif. TIF1 contains a single motif (residues 722-732) within the minimal region known to be required for its interaction with NRs [11, 25]. The truncated proteins TRIPs2-5, TRIP8 and TRIP9, which were isolated in a two hybrid screen for TR-interacting proteins [12], each contain at least one copy of the LXXLL motif (Fig. 3a), whereas the motif was absent in TRIPs whose interaction with TR was ligand-independent. An alignment of a selection of these sequences is shown in Fig. 3a, while Fig. 3b shows the incidence of motifs in the sequences of RIP140, SRC1a, TIF1, TIF2, CBP and p300, and the

boundaries of known receptor interaction domains in these proteins. Interestingly, motifs were also identified in several other proteins for which evidence exists of interaction with NRs, including Ara70 [26], SW13 [27], and the RelA (p65) subunit of NFκ-B [28], although the receptor interaction domains in these proteins have not been mapped. The ability of other proteins containing LXXLL motifs to bind to NRs will depend on their subcellular localisation, as well as the α-helicity and surface accessibility of the motifs. While it is clear that the conserved leucine residues are essential for the function of the motif, other amino acids may also be important given the degree of sequence conservation of equivalent motifs in SRC1/TIF2 or CBP/p300.

As many NR binding proteins contain multiple copies of the LXXLL motif it remains to be established whether this facilitates the simultaneous contact of individual partners in homo- and heterodimers of NRs, or whether it serves to provide alternative interaction surfaces to accomodate conformational changes imposed by the binding of NRs to different response elements. The systematic mutation of LXXLL motifs in coactivators such as SRC1 and CBP may allow us to decouple crosstalk or synergy between different signal transduction pathways, and thus provide a better understanding of their proposed roles as coactivators and integrators.

Example 5

Two Hybrid Interaction Assays

The yeast reporter strain used for all two hybrid assays was W303-1B (HMLα MATα HMRa his3-11, 15 trp1-1 ade2-1 can1-100 leu2-3, 11, ura3) carrying the plasmid pRLΔ21-U3ERE which contains a lacZ reporter gene driven by three estrogen response elements (EREs) [29]. The plasmids pBL1 and pASV3 which express the human ER DNA binding domain (DBD) and the VP16 acidic activation domain (AAD) respectively [30], were used to generate DBD or AAD fusion proteins for two hybrid interaction analyses. DBD-LXXLL motif fusion proteins were generated by ligation of phosphorylated, annealed oligonucleotide pairs into the pBL1 vector. AAD-ER was constructed by cloning a PCR fragment encoding amino acids 282-595 of the human ER into pASV3. AAD-ER Mut was constructed in a similar fashion except that the amino acids M543 and L544 or ER were mutated to alanines by recombinant PCR. All fusion constructs were fully sequenced. Transformants containing

the desired plasmids were obtained by selection for the appropriate plasmid markers and were grown to late log phase in 15 ml of selective medium (yeast nitrogen base containing 1% glucose and appropriate supplements) in the presence or absence of 10^{-7}M 17-β-estradiol (E2).

The expression of DBD- and AAD- fusion proteins in yeast cell-free extracts was verified by immunodetection using a monoclonal antibody recognising the human ER (a gift from P. Chambon, Strasbourg). The antibody recognises the "F" region of the LBD in the human ER, and also the "F" region tag at the N-termini of the DBD fusion proteins [30]. Equal amounts of protein were electrophoresed on polyacrylamide gels and transferred to nitrocellulose for western blotting. The preparation of cell-free extracts by the glass bead method and the measurement of β-galactosidase activity in the extracts were performed as previously described [29]. Two hybrid experiments were repeated several times, and the data shown in Figs. 1a and 1b represent reporter activities as measured in a single representative experiment. The β-galactosidase activities are expressed as nmoles/minute/μg protein.

Example 6

***In Vitro* Binding And Peptide Inhibition Assays**

GST-AF2 consists of the ligand binding domain of the mouse ER (amino acids 313-599) fused to glutathione-S-transferase and has been described previously [31]. GST-CBP consists of GST fused to the SRC1-binding domain of CBP and was constructed by cloning a PCR fragment encoding residues 2058-2163 of mouse CBP into the vector pGEX2TK (Pharmacia). Human SRC1a and SRC1e cDNAs were isolated form a human B cell cDNA library and cloned into a modified version of the expression vector pSG5. SRC1a M1234 and SRC1eM123 were constructed by recombinant PCR to introduce the mutations [L636A, L637A, L693A, L694A, L752A, L753A, L1438A, L1439A] or [L636A, L637A, L693A, L694A, L752A, L753A] respectively. All SRC1 constructs were fully sequenced. GST-SEPHAROSE™ beads were loaded with GST alone or GST-fusion proteins prepared from bacterial cell-free extracts. [^{35}S]-labelled SRC1 proteins were generated by *in vitro* translation and tested for interaction with GST proteins in the presence or absence of 10^{-6}M estradiol (E2) as previously described [21]. Binding was carried out for 3 hours at 4°C with gentle mixing in NETN buffer (100 mM NaCl, 1 mM EDTA, 0.5 % NP-40, 20 mM Tris HCl,

pH 8.0) containing protease inhibitors in a final volume of 1 ml. Peptides P-1 and P-2 were dissolved in water at a concentration of 4mg/ml and added individually to GST-binding reactions immediately before the addition of ligand. The increasing amounts of peptide added in the competition experiments shown corresponded to 2.5, 5, 12.5 and 25 μM.

Example 7

Transient Reporter Assays

Hela cells were transfected with 1μg of reporter 2ERE-pS2-CAT [32], 150 ng of β-galactosidase expression plasmid (internal control), 10 ng of ER expression plasmid and 50 or 200 ng of SRC1 expression plasmids or empty vector per well (in duplicate) using 24-well plates. Transfected cells were incubated overnight in Dulbecco's modified Eagle's medium without phenol red and containing 10% charcoal-treated FBS, and washed in fresh medium before addition of ligand (10^{-8}M E2) or vehicle. After 40 hrs, cells were harvested and extracts analysed for CAT and β-galactosidase activities [14, 21]. β-galactosidase activities were used to correct for differences in transfection efficiency.

Example 8

Pharmaceutical Composition

The following illustrates a representative pharmaceutical dosage form containing a peptide inhibitor and which may be used for therapy.

Injectable solution

A sterile aqueous solution, for injection, containing per ml of solution:

Peptide P-1	5.0mg
Sodium acetate trihydrate	6.8mg
Sodium chloride	7.2mg
Tween 20	0.05mg

A typical dose of peptide for adult humans is 30mg.

REFERENCES

1. Beato, M., Herrlich, P. & Schutz, G. *Cell* **83**, 851-857 (1995).
2. Mangelsdorf, D.J., *et al. Cell* **83**, 835-839 (1995)
3. Onate, S.A., Tsai, S.Y., Tsai, M.-J. & O'Malley, B.W. *Science* **270**, 1354-1357 (1995).
4. Kamei, Y., *et al. Cell* **85**, 403-414 (1996).
5. Yao, T.-P., Ku, G., Zhou, N., Scully, R. & Livingston, D.M. *Proc Natl Acad Sci USA* **93**, 10626-10631 (1996).
6. Voegel, J.J., Heine, M.J.S., Zechel, C., Chambon, P. & Gronemeyer, H. *EMBO J* **15**, 101-108 (1996).
7. Hong, H., Kohli, K., Trivedi, A., Johnson, D.L. & Stallcup, M.R. *Proc Natl Acad Sci USA* **93**, 4948-4952 (1996).
8. Chakravarti, D., *et al. Nature* **383**, 99-103 (1996).
9. Hanstein, B., *et al. Proc Natl Acad Sci USA* **93**, 11540-11545 (1996).
10. Cavailles, V., *et al. EMBO J* **14**, 3741-3751 (1995).
11. Le Douarin, B., *et al. EMBO J* **14**, 2020-2033 (1995).
12. Lee, J.W., Ryan, F., Swaffield, J.C., Johnston, S.A. & Moore, D.D. *Nature* **374**, 91-94 (1995).
13. vom Baur, E., *et al. EMBO J* **15**, 119-124 (1996).
14. Danielian, P.S., White, R., Lees, J.A. & Parker, M.G. *EMBO J* **11**, 1025-1033 (1992).
15. Bourguet, W., Ruff, M., Chambon, P., Gronemever, H. & Moras, D. *Nature* **375**, 377-3 82 (1995).
16. Renaud, J.-P., *et al. Nature* **378**, 681-689 (1995).
17. Wagner, R.L., *et al. Nature* **378**, 690-697 (1995).
18. Barettino, D., Ruiz, M.D.M.V. & Stunnenberg, H.G. *EMBO J* **13**, 3039-3049 (1994).
19. Durand, B., et al. *EMBO J* **13**, 5370-5382 (1994).
20. Saatcioglu, F., Bartunek, P., Deng, T., Zenke, M. & Karin, M. *Mol. Cell. Biol.* **13**, 3675-3685 (1993).
21. L'Horset, F., Dauvois, S., Heery, D.M., Cavailles, V. & Parker, M.G. *Mol. Cell. Biol.* **16**, 6029-6036 (1996).
22. Rost, B. & Sander, C. *Proc Natl Acad Sci USA* **90**, 7558-7562 (1993).
23. Kwok, R.P.S., *et al. Nature* **370**, 223-226 (1994).

24. Arias, J., *et al. Nature* **370**, 226-229 (1994).

25. Le Douarin, B., *et al. EMBO J* **15**, 6701-6715 (1996).

26. Yeh, S. & Chang, C. *Proc Natl Acad Sci USA* **93**, 5517-5521 (1996).

27. Yoshinaga, S.K., Peterson, C.L., Herskowitz, I. & Yamamoto, K.R. *Science* **258**, 1598- 1604 (1992).

28. Stein, B. & Yang, M.X. *Mol & Cell Biology* **15**, 4971-4979 (1995).

29. Metzger, D., Losson, R., Bornert, J.-M., Lemoine, Y. & Chambon, P. *Nucl Acid Res* **20**, 2813-2817 (1992).

30. Le Douarin, B., Pierrat, B., vom Baur, E., Chambon, P. & Losson, R. *Nucl Acid Res* **23**, 876-878 (1995).

31. Cavailles, V., Dauvois, S., Danielian, P.S. & Parker, M.G. *Proc Natl Acad Sci USA* **91**, 10009-10013 (1994).

32. Montano, M.M., Ekena, K., Krueger, K.D., Keller, A.L. & Katzenellenbogen, B.S. *Mol. Endocrinol* **10**, 230 -242 (1996).

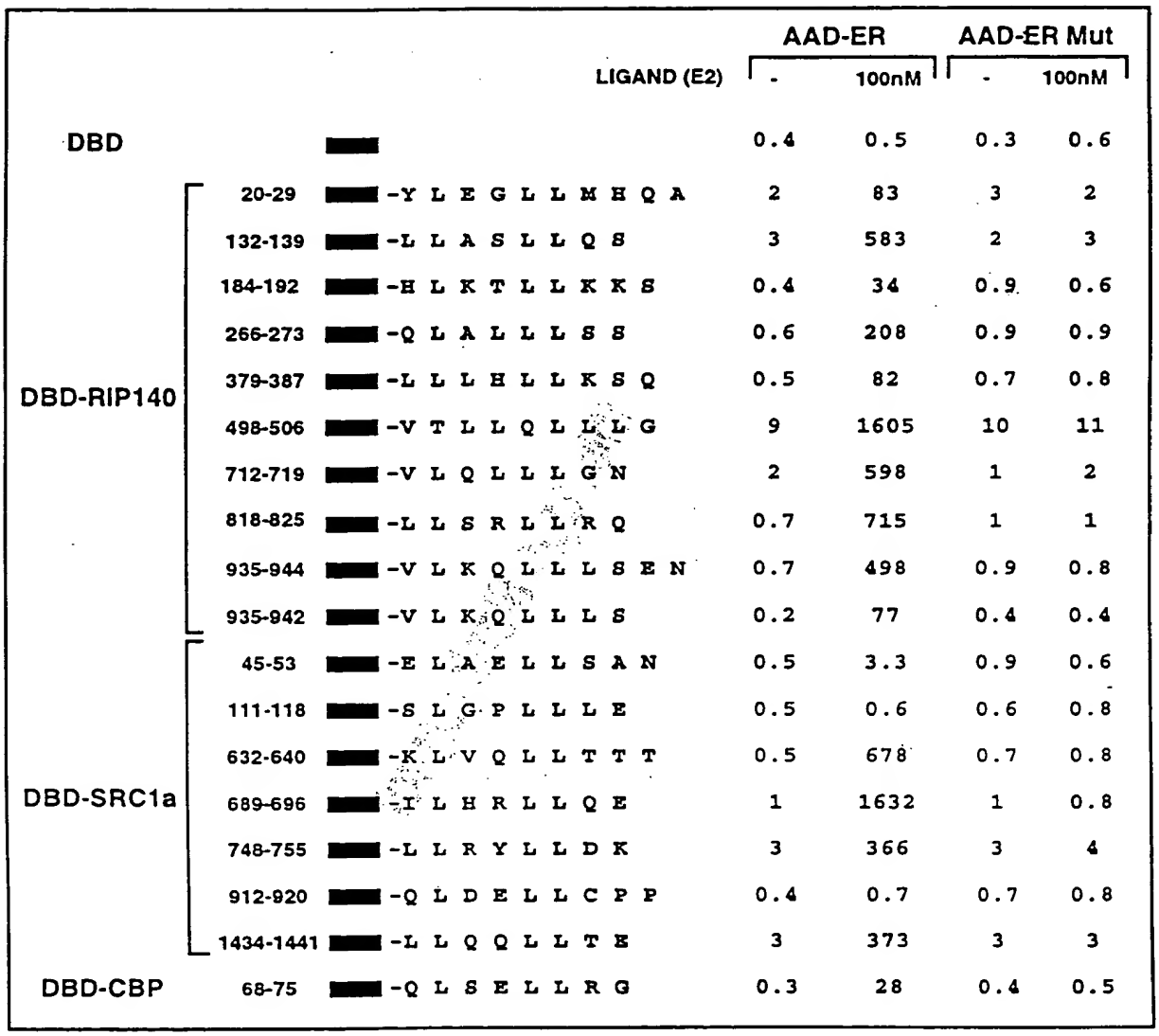
33. Eckner, R., *et al. Genes & Dev* **8**, 869-884 (1994).

EP97/021

AFG/MB : 29APR97

This Page Blank (uspto)

Figure 1A


AAD-ER
AAD-ER Mut
LIGAND (E2)
- 100nM - 100nM
DBD 0.4 0.5 0.3 0.6
DBD-RIP140
20-29 -Y L E G L L M H Q A 2 83 3 2
132-139 -L L A S L L Q S 3 583 2 3
184-192 -H L K T L L K K S 0.4 34 0.9 0.6
266-273 -Q L A L L L S S 0.6 208 0.9 0.9
379-387 -L L L H L L K S Q 0.5 82 0.7 0.8
498-506 -V T L L Q L L L G 9 1605 10 11
712-719 -V L Q L L L G N 2 598 1 2
818-825 -L L S R L L R Q 0.7 715 1 1
935-944 -V L K Q L L L S E N 0.7 498 0.9 0.8
935-942 -V L K Q L L L S 0.2 77 0.4 0.4
DBD-SRC1a
45-53 -E L A E L L S A N 0.5 3.3 0.9 0.6
111-118 -S L G P L L L E 0.5 0.6 0.6 0.8
632-640 -K L V Q L L T T T 0.5 678 0.7 0.8
689-696 -I L H R L L Q E 1 1632 1 0.8
748-755 -L L R Y L L D K 3 366 3 4
912-920 -Q L D E L L C P P 0.4 0.7 0.7 0.8
1434-1441 -L L Q Q L L T E 3 373 3 3
DBD-CBP
68-75 -Q L S E L L R G 0.3 28 0.4 0.5

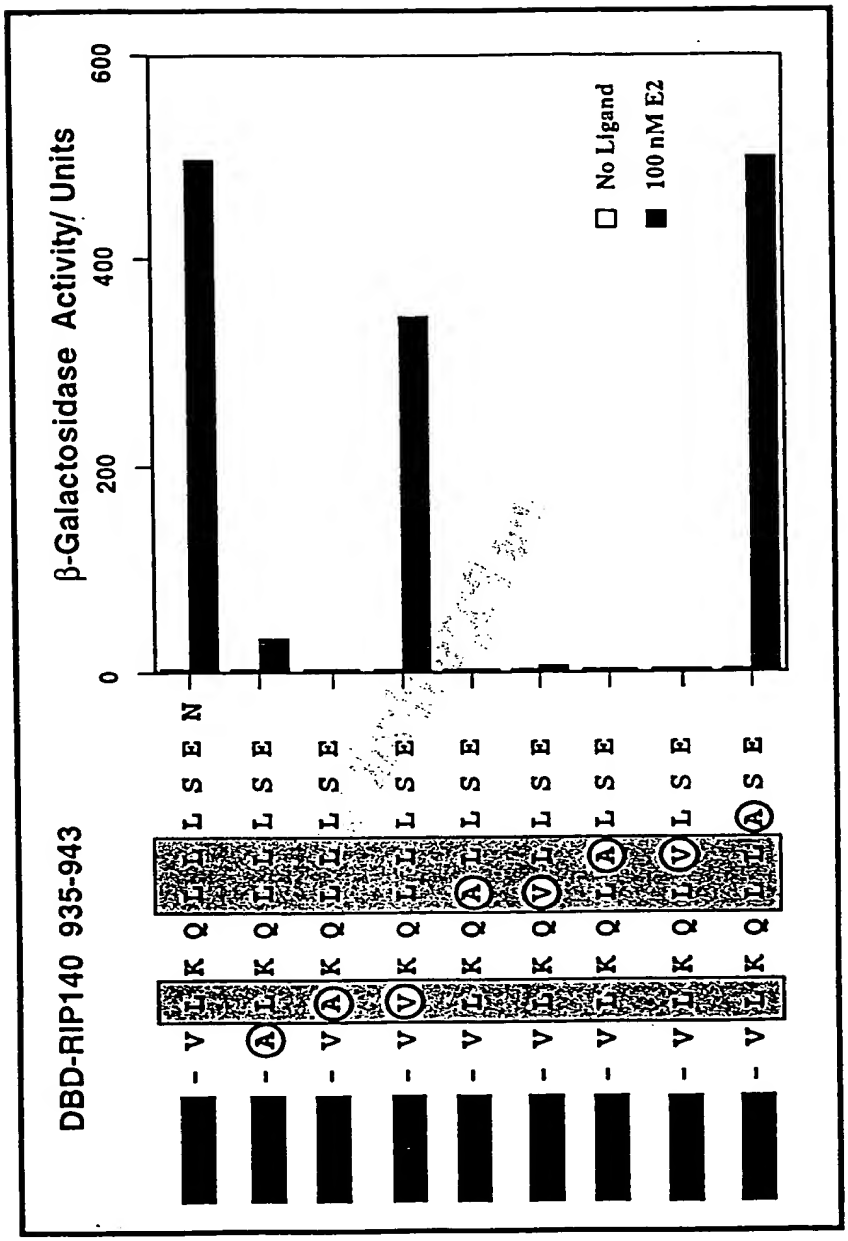

DBD-RIP140 935-943
β-Galactosidase Activity/ Units
0
200
400
600
No Ligand
100 nM E2
- V L K Q L L S E N
- A L K Q L L S E
- V A K Q L L S E
- V V K Q L L S E
- V L K Q A L S E
- V L K Q V L S E
- V L K Q L A S E
- V L K Q V L S E
- V L K Q L A S E
- V L K Q L L S E

Figure 1B

Figure 2A

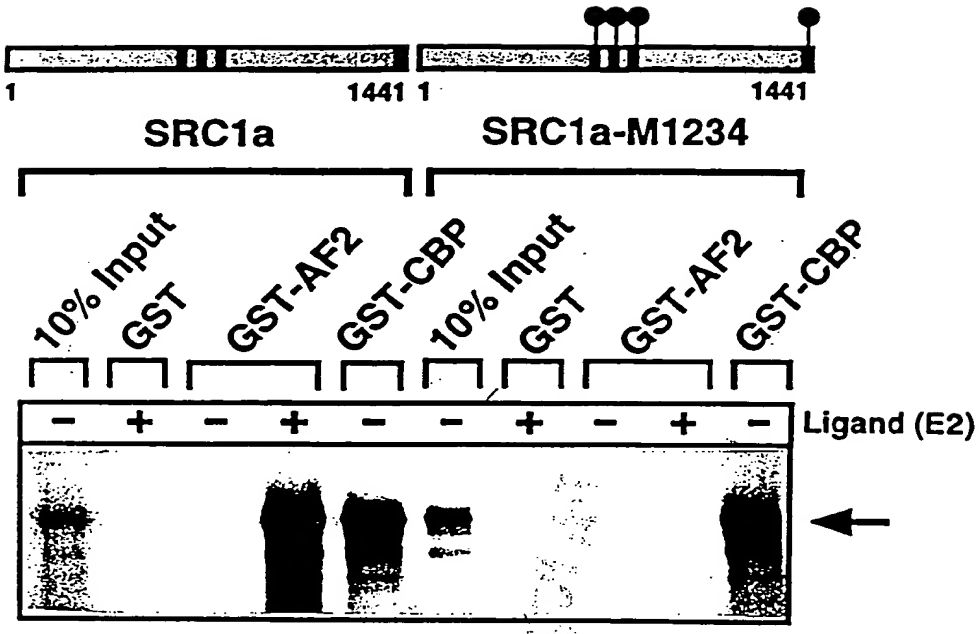


Figure 2B

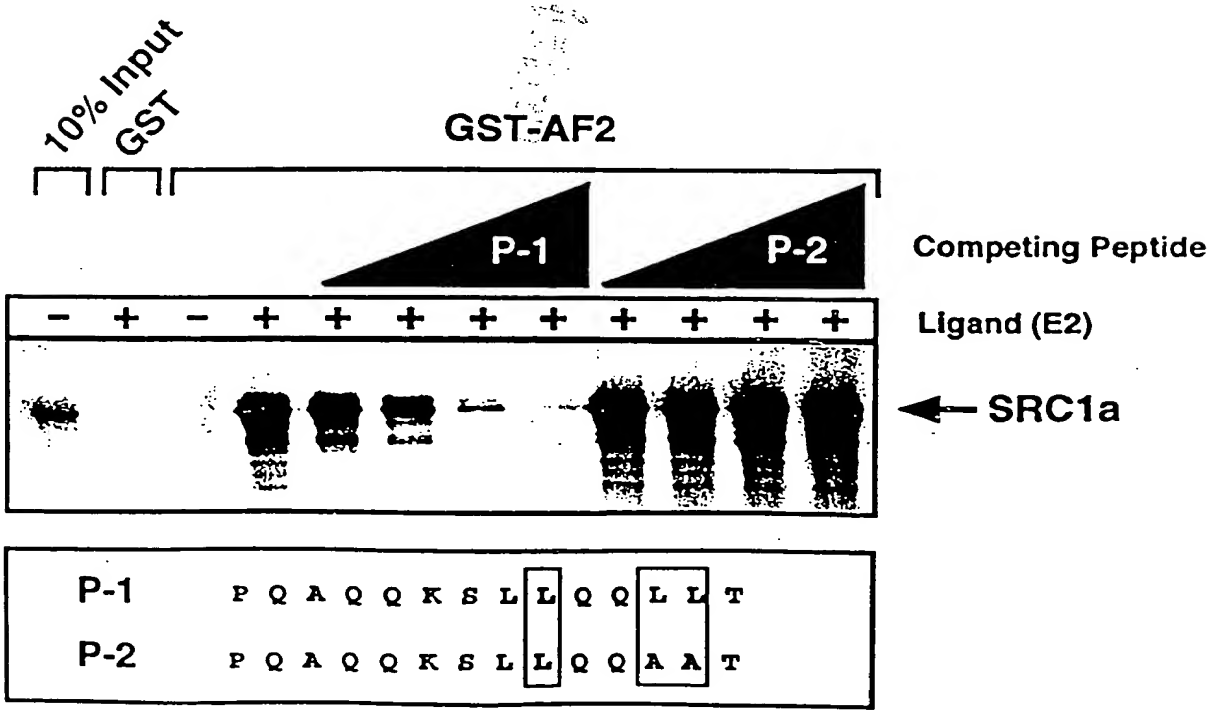


Figure 2C

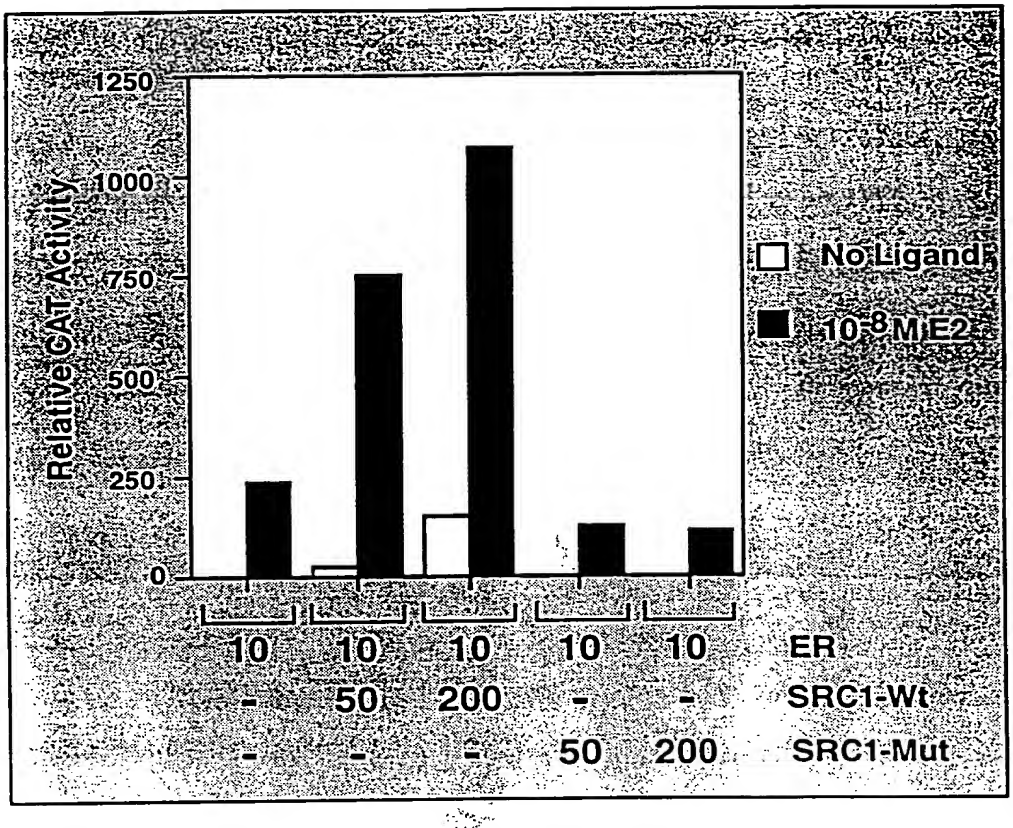
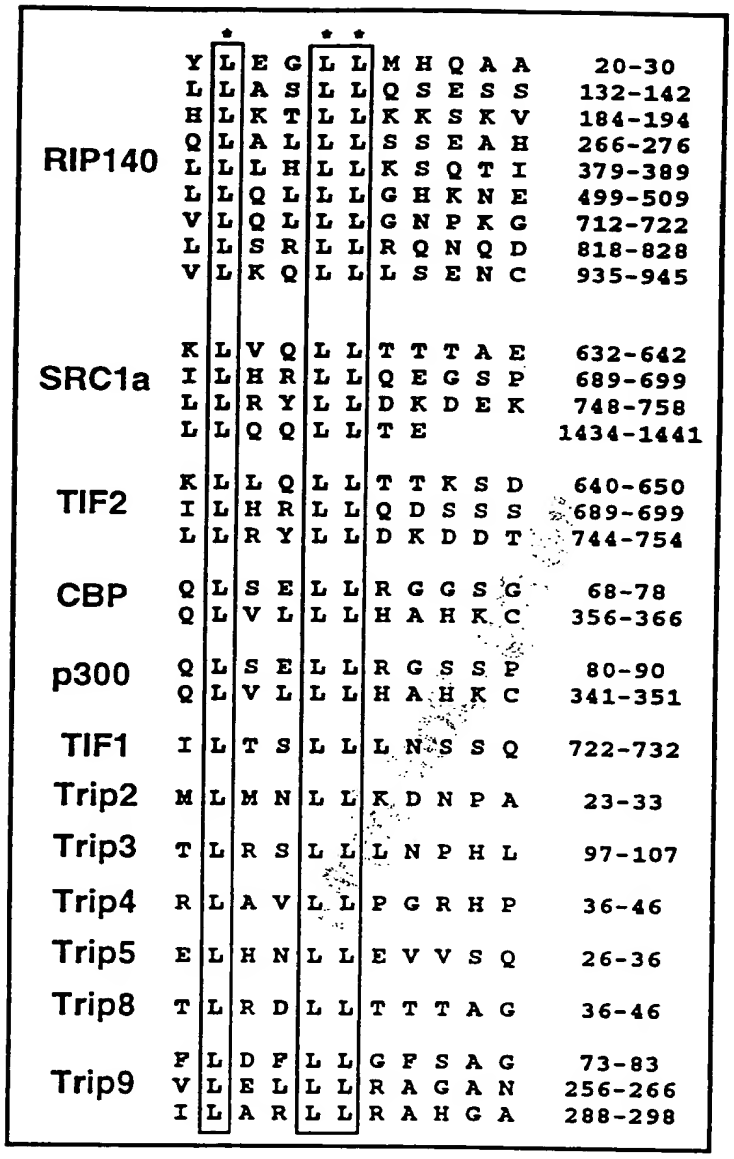

Relative CAT Activity
1250
1000
750
500
250
0
No Ligand
10-8M E2
10 10 10 10 10 ER
- 50 200 - - SRC1-Wt
- - - 50 200 SRC1-Mut

Figure 3A


RIP140
Y L E G L L M H Q A A 20-30
L L A S L L Q S E S S 132-142
H L K T L L K K S K V 184-194
Q L A L L L S S E A H 266-276
L L L H L L K S Q T I 379-389
L L Q L L L G H K N E 499-509
V L Q L L L G N P K G 712-722
L L S R L L R Q N Q D 818-828
V L K Q L L L S E N C 935-945
SRC1a
K L V Q L L T T T A E 632-642
I L H R L L Q E G S P 689-699
L L R Y L L D K D E K 748-758
L L Q Q L L T E 1434-1441
TIF2
K L L Q L L T T K S D 640-650
I L H R L L Q D S S S 689-699
L L R Y L L D K D D T 744-754
CBP
Q L S E L L R G G S G 68-78
Q L V L L L H A H K C 356-366
p300
Q L S E L L R G S S P 80-90
Q L V L L L H A H K C 341-351
TIF1
I L T S L L L N S S Q 722-732
Trip2
M L M N L L K D N P A 23-33
Trip3
T L R S L L L N P H L 97-107
Trip4
R L A V L L P G R H P 36-46
Trip5
E L H N L L E V V S Q 26-36
Trip8
T L R D L L T T T A G 36-46
Trip9
F L D F L L G F S A G 73-83
V L E L L L R A G A N 256-266
I L A R L L R A H G A 288-298

This Page Blank (uspto)

Figure 3B